DOCUMENT RESUME

ED 050 177 24 TM 000 581

AUTHOR Amer, J. P.; Dyer, J. S.
TITLE A Decision Model for Evaluating Potential Change in Instructional Programs.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY Department of Health, Education, and Welfare, Washington, D.C. Office of the Commissioner of Education.
REPORT NO CSE-62
BUREAU NO BR-6-1646
PUB DATE Nov 70
CONTRACT OEC-4-6-061646-1909
NOTE 31p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Cost Effectiveness, *Decision Making, Educational Improvement, Elementary Schools, Estimated Costs, Instructional Programs, *Models, Preschool Education, *Program Budgeting, Program Development, *Program Evaluation, *School Administration, Secondary Schools, Statistical Analysis

ABSTRACT

A statistical model designed to assist elementary school principals in the process of selection educational areas which should receive additional emphasis is presented. For each educational area, the model produces an index number which represents the expected "value" per dollar spent on an instructional program appropriate for strengthening that area. Although the model is explained in terms of this specific application, the approach proposed could also be used to structure similar decision problems at the district or state levels, or in secondary and pre-school educational systems. The calculation of the index number for a particular area depends on: (1) the relative importance of that area in terms of the overall educational goals; (2) the "utility" or "value" to the decision maker of making an improvement in that area, given the current level of performance; (3) the probability distribution of the results of implementing a particular type of improvement program for that area, given the current level of performance; and (4) the cost of the program. These factors are combined into a statistical formula yielding the desired index number. The use of these indices as decision instruments is also discussed. Statistical data and a bibliography are included. (Author/AE)

ED050177

BR 6-1646
PA 24
EA

CENTER FOR THE STUDY OF EVALUATION
UCLA Graduate School of Education
Los Angeles, California

A DECISION MODEL FOR EVALUATING POTENTIAL CHANGE IN INSTRUCTIONAL PROGRAMS

J. P. Amor and J. S. Dyer

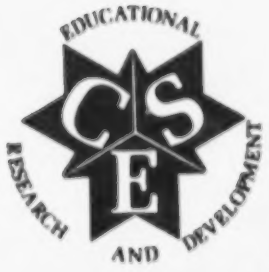
CSE Report No. 62
November 1970

TM 000 581

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

1

CENTER FOR THE STUDY OF EVALUATION



Marvin C. Alkin, Director

[redacted]

UCLA Graduate School of Education

The CENTER FOR THE STUDY OF EVALUATION is one of nine centers for educational research and development sponsored by the United States Department of Health, Education and Welfare, Office of Education. The research and development reported herein was performed pursuant to a contract with the U.S.O.E. under the provisions of the Cooperative Research Program.

Established at UCLA in June, 1966, CSE is devoted exclusively to finding new theories and methods of analyzing educational systems and programs and gauging their effects.

The Center serves its unique functions with an inter-disciplinary staff whose specialties combine for a broad, versatile approach to the complex problems of evaluation. Study projects are conducted in three major program areas: Evaluation of Instructional Programs, Evaluation of Educational Systems, and Evaluation Theory and Methodology.

This publication is one of many produced by the Center toward its goals. Information on CSE and its publications may be obtained by writing:

Office of Dissemination
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

This paper presents a model designed to assist elementary school principals in the process of selecting educational areas which should receive additional emphasis. For each educational area, the model produces an index number which represents the expected "value" per dollar spent on an instructional program appropriate for strengthening that area. Although the model will be explained in terms of this specific application, the approach proposed in this paper could also be used to structure similar decision problems at the district or state levels, or in secondary and pre-school educational systems.

The calculation of the index number for a particular area depends on the following factors: (1) the relative importance of that area; (2) the "utility," or "value" to the decision maker, of making an improvement in that area, given the current level of performance; (3) the probability distribution of the results of implementing a particular type of improvement program for that area, given the current level of performance; and (4) the cost of the above mentioned program. The first two factors will be discussed in the first section of this paper. In the second section the last two will be considered, and all four will be combined in a formula yielding the desired index number. The use of these indices as an aid to decision making will then be explained.

THE ESTIMATION OF UTILITY

This section describes how the "utility"* (of the decision maker) for the current state of the system is estimated. In our particular application, the system is an elementary school, the decision maker is its principal, and the state of the system is represented by the level of educational achievement of the school. As we shall see, the process of estimating utility presupposes the existence of (1) a well defined hierarchy of system objectives, and (2) adequate devices for measuring the degree of achievement of these objectives. We are then left with the problem of transforming performance measurements into a single (utility) number; this number reflects the decision maker's "satisfaction" with the state of the system as represented by these performance measures.

The formal structuring of a decision problem (in our case, the selection of educational areas to be emphasized or strengthened) must begin with the statement of an overall objective -- perhaps one as vague as "promote the good life". By asking <u>how</u> a given system contributes to the achievement of this "meta-objective", a hierarchy of primary objectives, secondary objectives, goals, and subgoals can be identified. For example, in elementary education, the primary objectives <u>might</u> be (1) the development of the child's personal or affective qualities, (2) the teaching of the basic skills required for further educational success, and (3) the introduction of the child to cultural norms

---

*For a discussion of the meaning of "utility", see the clasic work by J. Von Neumann and O. Morgenstern (1953) or that of Schlaifer (1959).

and values. Each of these three primary objectives can be subdivided into the secondary objectives which they logically and operationally imply. For example, objective (2) above may be subdivided into the teaching of verbal skills and the teaching of analytical skills. In a similar manner, secondary objectives may be subdivided into goals, and the latter into subgoals, etc. This subdivision process should continue until each "terminal" (lowest level) objective/goal in the hierarchy can be associated with an appropriate device for measuring the system's performance in that area. Figure 1 illustrates an example of the results of applying this objective subdivision process to elementary education. The numbers in parentheses refer to the terminal goal areas (in this case, subgoals) listed in Appendix I. For a further discussion of this process, as well as of some important caveats, see Pardee (1969), especially page 405.

The performance of an educational system is not usually determined by a direct measurement of how well it achieves its primary or secondary objectives. These objectives are too broad in nature for the development of valid and reliable measuring instruments. Instead, the state of the system may be assessed indirectly by measuring its performance in each terminal goal area for which adequate measuring devices exist. The Center for the Study of Evaluation at UCLA has investigated the existence of test instruments for each of the attributes listed in Appendix I (see Elementary School Evaluation Project, Booklet III). The results of this study indicate that numerous instruments are avail-

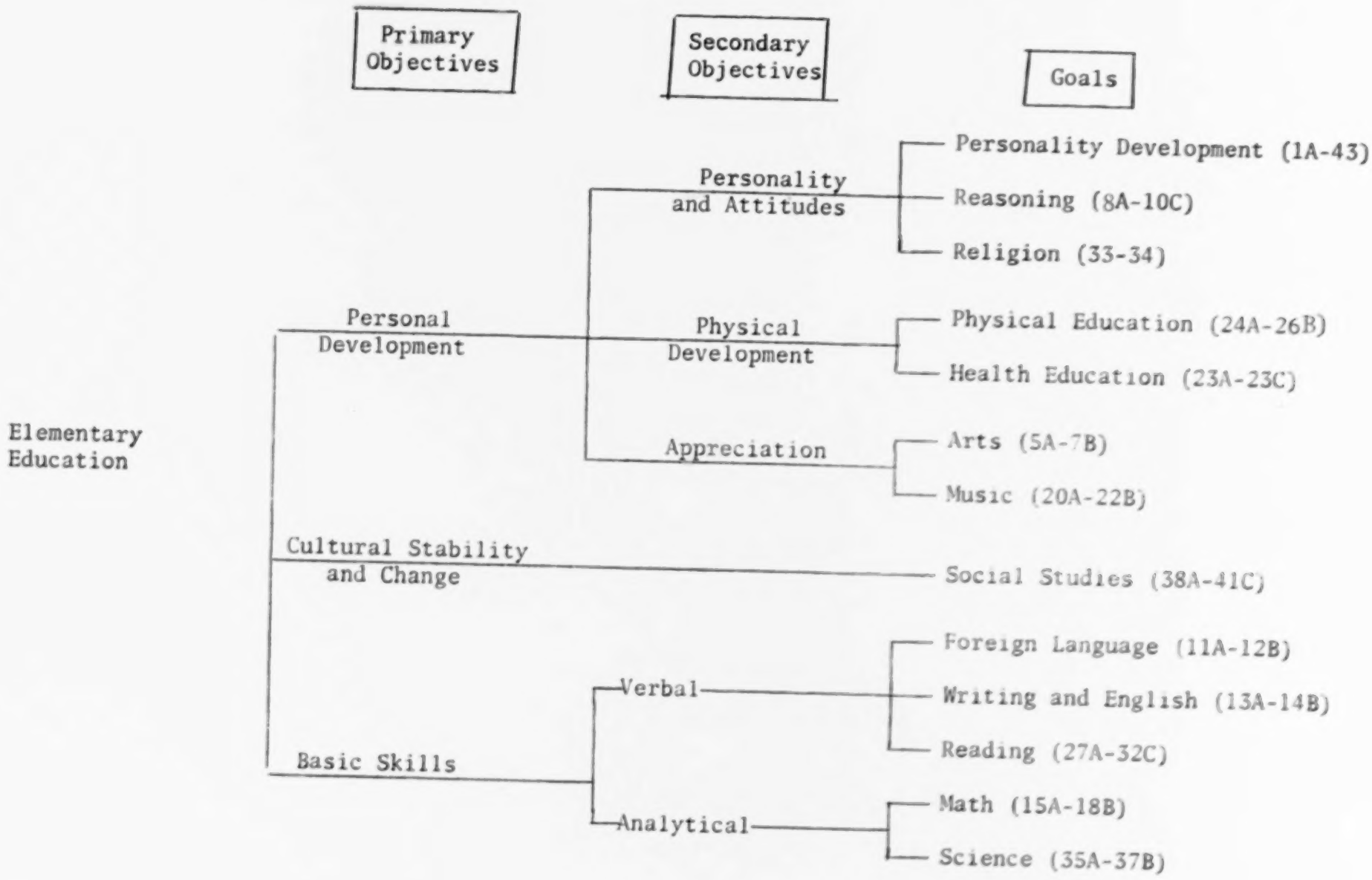


Figure 1. Illustrative Hierarchy of Objectives

able in the skill and cognitive areas, with their number and validity decreasing in the affective areas. However, continuing improvement in the number and validity of tests in all areas can be expected.

The results achieved on the various test instruments are generally expressed in terms of percentile scores or rankings on a national basis. Consequently, there is no absolute, invariant scale against which to measure performance, and we must take these scores to be our "raw" system performance measurements. It seems reasonable to assume that the "worth" or "value" (to the principal) of a given percentile score depends strongly on both (1) the particular goal area involved and (2) his aspiration level for that area. Since past achievements of a school depend to some extent on such exogenous input factors as the socioeconomic status of the parents, the location of the school (urban vs. rural), the region of the country, etc., one can reasonably expect these factors to influence the principal's aspiration level (and hence his utility for performance measurements) in each goal area.* So that the model will be sensitive to these environmental factors, performance data is currently being collected on schools' categorized according to environmental characteristics (see Elementary School Evaluation Project, Booklet IV).

*This will be discussed in a forthcoming Technical Report on the possible impact of such factors on the shape of the utility function for a given goal area.

If we accept the (percentile) scores obtained on standardized tests as adequate measures of the school's performance in the associated goal areas, it remains to be determined if these scores can be used directly in estimating the principal's utility. The contents of the previous paragraph and the following observations suggest that they cannot; that is, the scores must be transformed before they can be used for that purpose:

1. Given results in a particular goal area and ignoring all other results, it is clear that a score of 80 is considered twice as "good" as a score of 40. Also, it may not be true that a score increase from 40 to 50 has the same "value" as an increase from 80 to 90.
2. The "worth" of an increase from, say, 40 to 50 percentile points in two different goal areas may not be the same, (i.e., the principal may not be indifferent between these two outcomes).

The three words, "worth", "good", and "value" are often used synonymously with the term utility. However, in the last two paragraphs, these words have been used to express the utility of the decision maker with respect to only one goal area, and not the system as a whole. A key assumption of this model is that the principal's utility for a set of n percentile scores (one for each goal area, thus characterizing the state of the educational system) is simply the sum of his utility* for each

*When these utilities are measured on appropriate scales.

of the individual scores; therefore, it is evident that our next task is to transform the area scores into their associated "area utility values".

The above assumption about the decision maker's utility is equivalent to saying that his utility function is additively separable*. If we let

$n \equiv$ the number of terminal goal areas;

$a_i \equiv$ the (percentile) score obtained on a standardized test appropriate for measuring performance in area i;

$f_i(\cdot) \equiv$ the principals' standard**utility function for area i, i.e., it transforms the score for area i (a_i) into a number between 0 and 1 ($f_i(a_i)$);

$w_i \equiv$ the "weighing factor" for area i. If we require that $\sum_{i=1}^{n} w_i = 1$, this weight expresses the relative importance of area i with respect to the whole set of areas. The number w_i can also be viewed as the proper "scaling" factor for the standard utility function $f_i(\cdot)$ such that the principal's utility for the score a_i is given by the scaled utility function $\tilde{f}_i(a_i) = w_i f_i(a_i)$; and

$f(a_1,\ldots,a_n) \equiv$ the principal's utility for the set of n scores

$$a_i;\ i=1,\cdots,n$$

then the additive utility assumption says that :

---

* See Appendix II.

** The utility of a score of "0" is 0; the utility of a score of "100" is 1.

(1) $f(a_1,a_2,\cdots,a_n) = \tilde{f}_1(a_1) + \tilde{f}_2(a_2) + \cdots + \tilde{f}_n(a_n)$

or equivalently:

(2) $f(a_1,a_2,\cdots,a_n) = w_1 f_1(a_1) + w_2 f_2(a_2) + \cdots + w_n f_n(a_n)$

To transform the area score, a_i, into its associated area utility value, $\tilde{f}_i(a_i)$, we now need to determine the constant w_i and the function $f_i(\cdot)$.

The values of the w_i's may be obtained by asking the decision maker to compare the relative importance of goals on the same level within the hierarchy. For example, according to figure 1, the importance of personal development would be compared against the importance of cultural values and basic skills; at a lower level in the hierarchy, similarly, the importance of mathematics would be compared against that of science. The final weight for a sub-goal area would be obtained by multiplying its relative weight within its level of the hierarchy by the relative weight of its associated goal at the next highest level, times the relative weight of the next highest level goal/objective, etc. More detailed procedures for obtaining these weights are given in Pardee (1969). In addition, the problem of obtaining and synthesizing different weights (for each goal area of interest) from different groups which interact with the system (e.g., parents, teachers and administrators) is also considered in this document*

The function $f_i(\cdot)$ can only be approximated through the anal-

*For an alternate procedure, see reference, Elementary school Evaluation Project, Booklet III. This procedure is illustrated in terms of the 106 subgoals listed in Appendix I.

ysis of empirical data.* However, a few statements can be made about its expected shape. It seems reasonable that a_i, defined in terms of a percentile score, will be considered to be of greater value as it increases; that is, if $a_i^1 > a_i^2$, then $f_i(a_i^1) > f_i(a_i^2)$. Consequently, we may assume that the function $f_i(\cdot)$ is monotonically increasing on the closed interval $[0, 100]$, where the numbers in that interval refer to percentile scores. Additionally, there appear to be two particular percentile scores on a standardized test which serve as aspiration levels for the school principal: (1) the national norm (50th percentile score) and (2) the norms for schools of a particular "type," as characterized by the various environmental factors discussed earlier. Interviews with principals and the other individuals associated with elementary school systems indicated that an increase (in percentile score) of a given amount from a point below the national norm is considered to be of significantly greater value than the same amount of increase from a point above the national norm. This suggests that the slope of $f_i(\cdot)$ is steeper at points below the 50th percentile score than at points above this score. A similar behavior may be "expected" with respect to the "environmental" norm; however, current data limitations prevent us from verifying this hypothesis. This prediction of a decreasing slope for the utility function as the percentile score increases is also consistent with the "law of diminishing marginal utility" which has beem empirically

*For a discussion of several alternative approaches and a complete bibliography, see Fishburn (1967). For a discussion of how this was done in this particular project, see a forthcoming Technical Report.

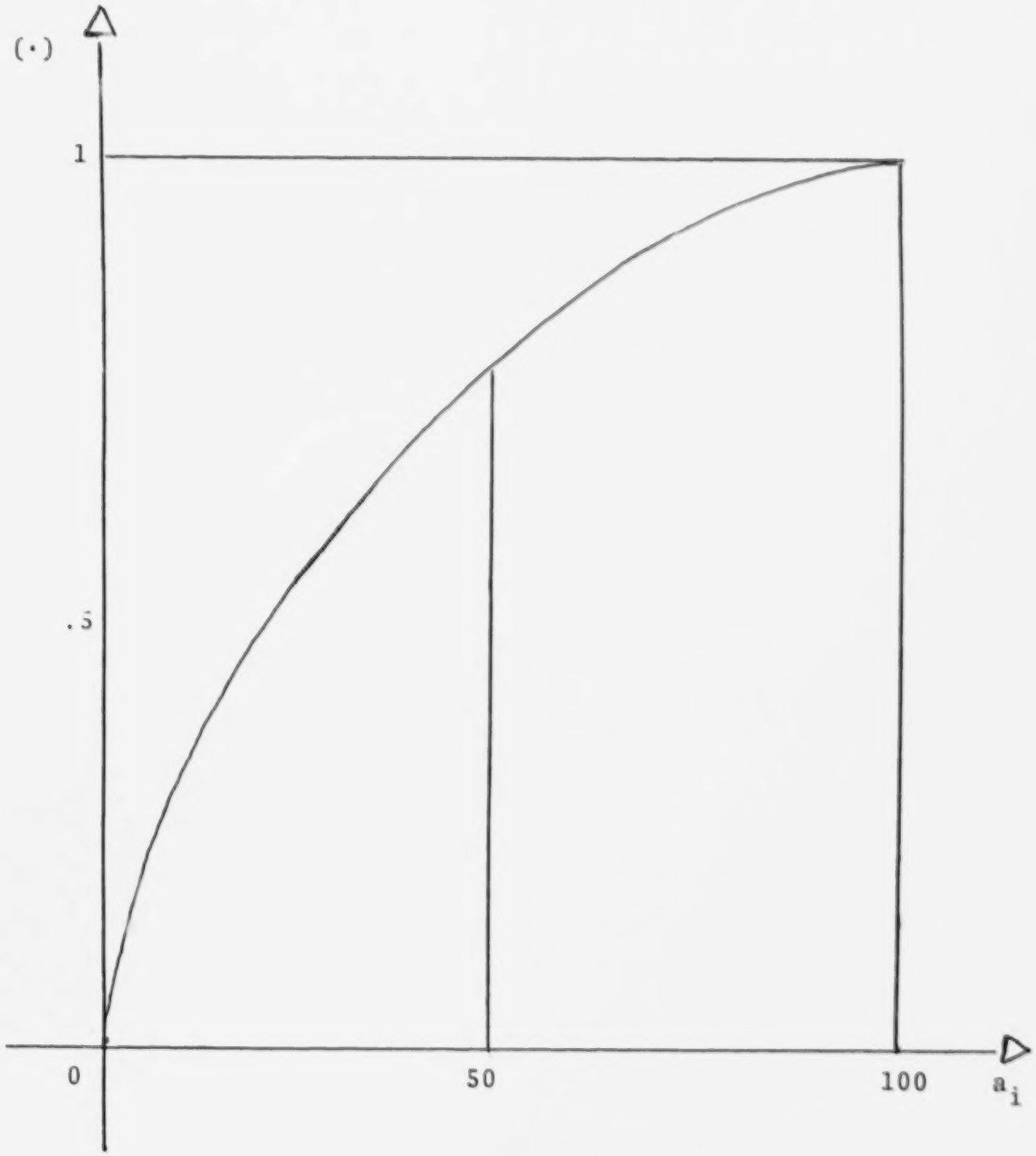


Figure 2 - A Possible Form of f_i (·)

verified in numerous studies in an economic context. One possible form of $f_i(\cdot)$ is shown in Fig. 2.

EXPECTED CHANGE IN UTILITY AS A DECISION CRITERION

This section describes how estimates of the decision maker's utility for various performance levels of the system may be combined with estimates of the effects on system performance of implementing various "programs" to provide a guide for decisions. In our particular application, the decision problem is to select educational goal areas which should receive more emphasis. The key assumption implicit in this discussion is that the decision maker prefers actions which maximize his expected utility.

It is reasonable to assume that a particular goal area will be selected for increased emphasis because (1) there exists an educational program(e.g., a new set of work books, the Sullivan reading program, a computer assisted arithmetic program, etc.), which has a "reasonably good chance" of improving the student's performance in that area, (2) a "significant" increase in the percentile score in that area results in a "significant" increase in the decision maker's utility, and (3) the cost of the program is within the school budget.

Because of the interaction among the exogenous input factors, it is impossible to state that "the adoption of a program of type j in area i will increase the performance from a_i^o to $a_i^o + \delta_i$." Instead, the potential results of adopting a program of type j in area i <u>should be</u> described by the conditional (or posterior) probability distribution of the scores which would be obtained upon retesting the students after implementation of the program.

Using this distribution, one could calculate the probability of achieving a specified result. We assume that the random variable a'_i (representing the score to be obtained upon retesting) has a probability density function which depends on 3 factors: (1) the particular subgoal area (i), (2) the current level of achievement (a_i°), and (3) the particular type of program selected (j), and we denote this conditional probability distribution by $g_i\ (a'_i \mid a_i^\circ, j)$. For example, the subgoal area may be "operations with integers," the current level of achievement may be "40th percentile ranking," and the type of program may be "a computer assisted arithmetic program." A possible form of $g_i\ (a'_i \mid 40, j)$ is shown in Figure 3.

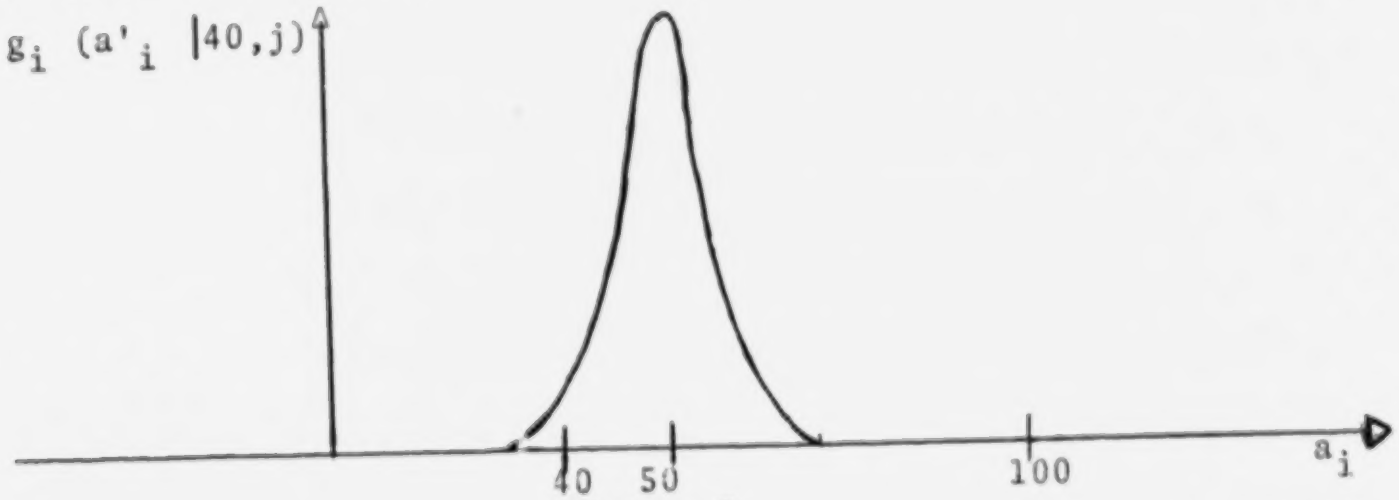


Figure 3
A Possible Form of a Posterior Probability Distribution of Scores

Unfortunately, it is extremely difficult to obtain objective information which would yield an estimate of this probability distribution. However, subjective information may be used to approximate the form of $g_i\ (a'_i \mid a_i^\circ, j)$. Persons who have observed the impact of new educational programs may be asked for their estimate of the probability of achieving a specific change in performance as a result of adopting a program of type j in area i, when the

current level of performance is a_i^o. For example, these "experts" may estimate the probability of an increase of from three to five percentile points in area i, given a_i^o and program j, to be .3. This implies that

$$(3) \qquad \int_{a_i^o+3}^{a_i^o+5} g_i(a'_i|a_i^o,j)da'_i \approx .3$$

We will denote a subjective estimate of the value of the integral of $g_i(a'_i|a_i^o,j)$ over a particular interval, Δ_k,* by $P_i(\Delta_k|a_i^o,j)$. By obtaining similar estimates for K non-overlapping intervals which cover the range of a'_i, we can obtain a discrete approximation to the desired distribution. By varying the size and number of the intervals we can make the approximation as "close" to the continuous function as we desire.** In addition, we naturally require that $\sum_{k=1}^{K} P_i(\Delta_k|a_i^o,j)=1$.

The theoretical probability distribution may be used to determine the expected change in utility, resulting from the implementation of program j in area i given that the current level of performance is a_i, as follows:

$$(4) \quad E(\Delta U_i) = w_i\left[\int_0^{100} f_i(a'_i)g_i(a'_i|a_i^o,j)da'_i - f_i(a_i^o)\right] ***$$

where E ≡ expected value operator,

ΔU_i ≡ change in utility in area i,

---

*Note that the symbol Δ_k carries information about both the location and the size of the interval of interest.

**See Schlaifer (1959) for a further discussion of this topic.

***See Appendix II for a theoretical discussion of the assumptions implicit in this formulation.

and the other variables are identical to those defined in the previous section. If the discrete approximation to $g_i(a'_i|a_i^\circ,j)$ is used, expression (4) simplifies to

$$(5)\quad E(\Delta U_i) \equiv w_i \left[\overline{\{\sum_{k=1}^{K} f_i(\Delta_k) \cdot P_i(\Delta_k|a_i^\circ,j)\}} - f_i(a_i^\circ) \right]$$

where the value of $\overline{f_i}(\Delta_k)$ may be approximated by the value of $f_i(\cdot)$ at the midpoint of the interval Δ_k.

Since it is anticipated that decisions must be made within a budget constraint, an additional modification of expressions (4) and (5) is required. Assume that the cost of a program of type j has been estimated to be c_j. Dividing these expressions by c_j would provide an estimate of the expected change in utility per dollar associated with a program of type j in area i, given a_i°. This figure represents a more desirable guide to action, especially if the programs in question are divisible.*

To summarize, the implications of expression (5) may be clarified if it is written in an equivalent form and divided by c_j as shown in (6).

$$(6)\quad \frac{E(\Delta U_i)}{c_j} = \frac{w_i \cdot \left[\overline{\sum_{k=1}^{K} \{f_i(\Delta_k) - f_i(a_i^\circ)\} \cdot P_i(\Delta_k | a_i^\circ,j)} \right]}{c_j}$$

The numerator is the product of the measure of relative importance, w_i, the change in utility associated with a given change

*By divisible, we mean that they may be adopted in part. Methods for dealing with this problem if divisibility is not a reasonable assumption will be considered in a forthcoming report.

in performance, $f_i(\Delta_k) - f_i(a_i^\circ)$, and the subjective probability of that change, $P_i(\Delta_k \mid a_i^\circ, j)$. If the probabilities that the resulting score will fall in each of several disjoint intervals Δ_k are estimated, their product with their associated measures of change in utility must be summed. In words, expression (6) for the expected change in utility per dollar is approximately equivalent to

$$\frac{\text{a measure of the importance of area i} \cdot \text{the change in utility associated with the adoption of a program of type j in area i} \cdot \text{the probability of achieving that change}}{\text{the cost of a program of type j}}$$

CONCLUSION

The model which was described in the preceding sections provides the decision maker (an elementary school principal) with index numbers which represent estimates of the expected changes in his utility for the performance of his school which would result from the adoption of particular types of programs in particular areas. It is felt that these index numbers would provide valuable information in his dealing with the problem of identifying areas in which action should be considered, and identifying the types of programs which would provide the greatest expected

contribution to the achievement of his instructional goals.

The actual use of the model by an elementary shcool principal could be encouraged through the provision of a series of tables containing the values

$$\frac{\sum_{k=1}^{K} \{f_i(\Delta_k) - f_i(a_i^\circ)\} \cdot P_i(\Delta_k | a_i, j)}{c_j}$$

for each area, i, for a series of scores, a_i°, and for several types of programs, j. The principal would merely be required to determine his own measures of the relative importance of each area (the w_i's), and obtain the remainder of the information directly from the tables.

REFERENCES

Elementary School Evaluation Project, Booklet II, Center for the Study of Evaluation, UCLA.

Elementary School Evaluation Project, Booklet III, Center for the Study of Evaluation, UCLA.

Elementary School Evaluation Project, Booklet IV, Center for the Study of Evaluation, UCLA.

Fishburn, P. C. Methods of estimating additive utilities. Management Science, 1967, 13 (7).

Miller, J. R., III. Assessing alternative transportation systems. RM-5865-DOT, The RAND Corporation, Santa Monica, California, April, 1969.

Pardee, F. S., et. al. Measurement and evaluation of transportation system effectiveness, RM-5869-DOT, The RAND Corporation, Santa Monica, California, October, 1969.

Schlaifer, R. Probability and statistics for business decisions. New York: McGraw-Hill, 1959.

Von Neuman, J., & Morgenstern, O. Theory of games and economic behavior. Princeton, N. J.: Princeton University Press, 1953.

APPENDIX I

OUTLINE OF 145 GOALS OF ELEMENTARY SCHOOL EDUCATION

AFFECTIVE
- 1. TEMPERAMENT: PERSONAL
  - A. Shyness-Boldness
  - B. Neuroticism-Adjustment
  - C. General Activity-Lethargy
- 2. TEMPERAMENT: SOCIAL
  - A. Dependence-Independence
  - B. Hostility-Friendliness
  - C. Socialization-Rebelliousness
- 3. ATTITUDES
  - A. School Orientation
  - B. Self Esteem
- 4. NEEDS AND INTERESTS
  - A. Need Achievement
  - B. Interest Areas

ARTS-CRAFTS
- 5. VALUING ARTS AND CRAFTS
  - A. Appreciation of Arts and Crafts
  - B. Involvement in Arts and Crafts
- 6. PRODUCING ARTS AND CRAFTS
  - A. Representational Skill in Arts and Crafts
  - B. Expressive Skill in Arts and Crafts
- 7. UNDERSTANDING ARTS AND CRAFTS
  - A. Arts and Crafts Comprehension
  - B. Developmental Understanding of Arts and Crafts

COGNITIVE
- 8. REASONING
  - A. Classificatory Reasoning
  - B. Relational-Implicational Reasoning
  - C. Systematic Reasoning
  - D. Spatial Reasoning
- 9. CREATIVITY
  - A. Creative Flexibility
  - B. Creative Fluency
- 10. MEMORY
  - A. Span and Serial Memory
  - B. Meaningful Memory
  - C. Spatial Memory

FOREIGN LANGUAGE
- 11. FOREIGN LANGUAGE SKILLS
  - A. Reading Comprehension of a Foreign Language
  - B. Oral Comprehension of a Foreign Language
  - C. Speaking Fluency in a Foreign Language
  - D. Writing Fluency in a Foreign Language

12. FOREIGN LANGUAGE ASSIMILATION
    A. Cultural Insight through a Foreign Language
    B. Interest in and Application of a Foreign Language

LANGUAGE ARTS

13. LANGUAGE CONSTRUCTION
    A. Spelling
    B. Punctuation
    C. Capitalization
    D. Grammar and Usage
    E. Penmanship
    F. Written Expression
    G. Independent Application of Writing Skills

14. REFERENCE SKILLS
    A. Use of Data Sources as Reference Skills
    B. Summarizing Information for Reference

MATHEMATICS

15. ARITHMETIC CONCEPTS
    A. Comprehension of Numbers and Sets in Mathematics
    B. Comprehension of Positional Notation in Mathematics
    C. Comprehension of Equations and Inequalities
    D. Comprehension of Number Principles

16. ARITHMETIC OPERATIONS
    A. Operations with Integers
    B. Operations with Fractions
    C. Operations with Decimals and Percents

17. MATHEMATICAL APPLICATIONS
    A. Mathematical Problem Solving
    B. Independent Application of Mathematical Skills

18. GEOMETRY
    A. Geometric Facility
    B. Geometric Vocabulary

19. MEASUREMENT
    A. Measurement Reading and Making
    B. Statistics

MUSIC

20. MUSIC APPRECIATION AND INTEREST
    A. Music Appreciation
    B. Music Interest and Enjoyment

21. MUSIC PERFORMANCE
    A. Singing
    B. Musical Instrument Playing
    C. Dance (Rhythmic Response)

22. MUSIC UNDERSTANDING
    A. Aural Identification of Music
    B. Music Knowledge

PHYSICAL EDUCATION - HEALTH - SAFETY

23. HEALTH AND SAFETY
    - A. Practicing Health and Safety Principles
    - B. Understanding Health and Safety Principles
    - C. Sex Education

24. PHYSICAL SKILLS
    - A. Muscle Control (Physical Education)
    - B. Physical Development and Well-Being (Physical Education)

25. SPORTSMANSHIP
    - A. Group Activity - Sportsmanship
    - B. Interest in and Independent Participation in Sports and Games

26. PHYSICAL EDUCATION
    - A. Understanding of Rules and Strategies of Sports and Games
    - B. Knowledge of Physical Education Apparatus and Equipment

READING

27. ORAL-AURAL SKILLS
    - A. Listening Reaction and Response
    - B. Speaking

28. WORD RECOGNITION
    - A. Phonetic Recognition
    - B. Structural Recognition

29. READING MECHANICS
    - A. Oral Reading
    - B. Silent Reading Efficiency

30. READING COMPREHENSION
    - A. Recognition of Word Meanings
    - B. Understanding Ideational Complexes
    - C. Remembering Information Read

31. READING INTERPRETATION
    - A. Inference Making from Reading Selections
    - B. Recognition of Literary Devices
    - C. Critical Reading

32. READING APPRECIATION AND RESPONSE
    - A. Attitude toward Reading
    - B. Attitude and Behavior Modification from Reading
    - C. Familiarity with Standard Children's Literature

RELIGION

33. RELIGIOUS KNOWLEDGE

34. RELIGIOUS BELIEF

SCIENCE

35. SCIENTIFIC PROCESSES
    - A. Observation and Description in Science
    - B. Use of Numbers and Measures in Science
    - C. Classification and Generalization in Science
    - D. Hypothesis Formation in Science

E. Operational Definitions in Science
F. Experimentation in Science
G. Formulation of Generalized Conclusions in Science

36. SCIENTIFIC KNOWLEDGE
A. Knowledge of Scientific Facts and Terminology
B. The Nature and Purpose of Science

37. SCIENTIFIC APPROACH
A. Science Interest and Appreciation
B. Application of Scientific Methods to Everyday Life

SOCIAL STUDIES

38. HISTORY AND CIVICS
A. Knowledge of History
B. Knowledge of Governments

39. GEOGRAPHY
A. Knowledge of Physical Geography
B. Knowledge of Socio-Economic Geography

40. SOCIOLOGY
A. Cultural Knowledge
B. Social Organization Knowledge

41. APPLICATION OF SOCIAL STUDIES
A. Research Skills in Social Studies
B. Citizenship
C. Interest in Social Studies

APPENDIX II

Let the state of the educational system be partially characterized by the n-tuple $(a_1,---,a_n)$ of percentile scores obtained on tests covering the n subgoal areas.* It is reasonable (and indeed necessary) to assume the existence of a cardinal utility function f: $R^n \rightarrow R^1$, which assigns to each possible state of the system a real number representing its "worth" to the decision maker. Let the initial state of the system be $a^\circ \equiv (a_1^\circ,---,a_n^\circ)$, with corresponding utility number $f(a^\circ)$, and suppose that an improvement program** is administered in each subgoal area, and then the students are retested. One would expect the resulting new set of scores $a'=(a_1',---,a_n')$ to dominate*** the old one and the new utility number $f(a')$ to be greater than $f(a^\circ)$. Whether or not this occurs depends on various uncontrollable factors, so that until the new scores are known, they and their associated utility number must be treated as random variables. A problem of interest is therefore to estimate, in advance, the change in the decision maker's utility if, given an initial state a°, improvement programs were administered and the students retested as discussed above. More importantly, we wish to estimate the "contribution" of each program (to the estimated change

---

*Clearly other system attributes should be included for a "complete" characterization.

**Any program applicable to a specified area will do. It could be the "best" program available for that area. For purposes of this discussion it matters only that the program be specific enough to have a cost.

***$a_i' \geq a_i^\circ$ for all i=1,---,n and $a_1' > a_1^\circ$ for at least one i.

in the decision maker's utility) and relate this contribution to the cost of that program in a way which will be meaningful to the decision maker in his attempt to maximize his expected utility by selecting costly improvement programs, while staying within his limited budget.

It is clear that the "utility contribution" of each program depends on various factors, three of which are its quality relative to that of the programs covering the other subgoal areas, the shape of the decision maker's utility function, and the initial state of the system. It is therefore no trivial matter to estimate these "utility contributions." However, if the utility function is additively separable* and if the improvement programs are independent,** then the expected change in utility is simply the sum of the expected utility contributions of the individual programs. For the sake of clarity and ease of presentation we shall show this for n=2, the case for larger values of n clearly follows.

For n=2 then, after the initial testing, we have a 2-tuple of percentile scores $a^\circ \equiv (a_1^\circ, a_2^\circ)$ and a utility number of $f(a^\circ) \equiv f(a_1^\circ, a_2^\circ)$. If certain improvement programs are contemplated for these two areas, the scores which might be obtained after retesting are now random variables $a' \equiv (a_1', a_2')$

---

*$f(a_1, a_2, \ldots, a_n) = \tilde{f}_1(a_1) + \tilde{f}_2(a_2) + \ldots + \tilde{f}_n(a_n)$; or equivalently with standard rescaling of the individual $\tilde{f}_i$'s: $f(a_1, \ldots a_n) = w_1 f_1(a_1) + w_2 f_2(a_2) + \ldots + w_n f_n(a_n)$.

**No spillover or interaction effects.

with joint conditional probability distribution $g(a_1', a_2'/a_1^\circ, a_2^\circ)$.* Clearly the utility for these new scores $f(a_1', a_2')$ is also a random variable, so that the expected utility of the decision maker for the new state of the system a', given the initial state a° and the improvement programs chosen, is given by:**

$$(1)\quad E\left[\overline{f(a_1', a_2'/a_1^\circ, a_2^\circ)}\right] = \iint f(a_1', a_2')\, g(a_1', a_2'/a_1^\circ, a_2^\circ)\, da_1' da_2'$$

and the expected change in utility is therefore

$$(2)\quad E[\Delta f] = E\left[\overline{f(a_1', a_2'/a_1^\circ, a_2^\circ)}\right] - f(a_1^\circ, a_2^\circ).$$

The assumption of program independence means that

$$(3)\quad g(a_1', a_2'/a_1^\circ, a_2^\circ) = g_1(a_1'/a_1^\circ)\, g_2(a_2'/a_2^\circ)$$

where g_1 and g_2 represent the conditional probability distribution*** of the random variables a_1' (given a_1°) and a_2' (given a_2°) respectively. Combining this assumption with that of separable utility and substituting in (1) yields

$$(4)\quad E\left[\overline{f(a_1', a_2'/a_1^\circ, a_2^\circ)}\right] = \iint \left[\overline{w_1 f_1(a_1') + w_2 f_2(a_2')}\right] g_1(a_1'/a_1^\circ)\, g_2(a_2'/a_2^\circ)\, da_1' da_2'$$

$$= w_1 \iint f_1(a_1')\, g_1(a_1'/a_1^\circ)\, g_2(a_2'/a_2^\circ)\, da_1' da_2' +$$

$$w_2 \iint f_2(a_2')\, g_1(a_1'/a_1^\circ)\, g_2(a_2'/a_2^\circ)\, da_1' da_2'$$

---

*This distribution is unknown to us at this time. In this Appendix we avoid complicating the expressions by not including indices which would identify the particular programs.

**Assume continuous functions here. The integrations are over the interval [0,100].

***These distributions are also unknown to us at this time.

$$= w_1 \int f_1(a_1')\, g_1(a_1'/a_1^\circ)\, da_1' + w_2 \int f_2(a_2')\, g_2(a_2'/a_2^\circ)\, da_2'$$

$$= w_1 E\left[f_1(a_1'/a_1^\circ)\right] + w_2 E\left[f_2(a_2'/a_2^\circ)\right]$$

Now substituting in (2) we get

$$(5)\quad E[\Delta f] = w_1 E\left[f_1(a_1'/a_1^\circ)\right] + w_2 E\left[f_2(a_2'/a_2^\circ)\right] - \left[w_1 f_1(a_1^\circ) + w_2 f_2(a_2^\circ)\right]$$

$$= w_1 \{E[f_1(a_1'/a_1^\circ)] - f_1(a_1^\circ)\} + w_2 \{E[f_2(a_2'/a_2^\circ)] - f_2(a_2^\circ)\}$$

$$= w_1 E[\Delta f_1] + w_2 E[\Delta f_2]$$

so that the expected "total" change in utility is just the sum of the expected utility changes for the subgoal areas. Since the assumption of program independence implies that the utility change in any subgoal area is solely* attributable to the improvement program selected for that area, we may take the expected change in utility for a given area as a measure of the expected value of that program to the decision maker (i.e., its expected utility contribution $= w_i E[\Delta f_i]$), given a certain performance level in that area.

The expected utility contribution of a program is the number we wish to relate to the cost of that program to assist the decision maker in the program selection process. In order to obtain this number we must estimate $E\left[f_i(a_1'/a_i^\circ)\right]$ ** (the expected, new, unweighed utility value in area i, given a_i°)

---

*Actually we only mean to say that programs covering other areas do not contribute to the utility change in this area. It is perhaps too strong to use the word "solely" here as there are surely other factors and random disturbances which affect this utility change.

**Recall from the main body of the paper that we have already determined a_1°, f_i and w_i; i=1,2.

for i=1,...,n, since, as mentioned earlier, the $g_i(a_i'/a_i^\circ)$* are unknown to us.

One approach in estimating $E\left[f_i(a_i'/a_i^\circ)\right]$ directly is to specify a certain number of percentile points δ_i and then to estimate, by questioning the appropriate "experts," the probability that the students will score $a_i^\circ+\delta_i$ percentile points or better, after being exposed to a particular improvement program, given an initial score of a_i°. Let us denote this probability by p_i and call it the <u>probability of success.</u> Then

$$p_i \equiv \text{Prob}\ (a_i' \geq a_i^\circ+\delta_i \mid a_i^\circ)$$

and an approximation to the expected new utility value in area i is given by

$$(6)\quad E\left[f_i(a_i'/a_i^\circ)\right] \approx p_i f_i(a_i^\circ+\delta_i) + (1-p_i)\ f_i(a_i^\circ)$$

In this approximation, the first term on the right hand side of (6) is a conservative estimate of $\int_{a_i^\circ+\delta_i}^{100} f_i(a_i')g_i(a_i'/a_i^\circ)da_i'$ since the utility function is evaluated at the single point of $a_i^\circ+\delta_i$, and we know that this function is monotonically increasing. Without knowledge of the parameters of $g_i(a_i'/a_i^\circ)$ it is not clear whether or not $(1-p_i)f_i(a_i^\circ)$ is a conservative estimate of $\int_0^{a_i^\circ+\delta_i} f_i(a_i')g_i(a_i'/a_i^\circ)da_i'$. However, if there are good reasons to believe that the variance of $g_i(a_i'/a_i^\circ)$ is relatively small, then the above approximation to $E\left[f_i(a_i'/a_i^\circ)\right]$

---

*It is conceivable that in the future there will be sufficient historical data to estimate this function empirically. Perhaps it can be done presently if we do not condition it on a_i°. However, for present purposes we shall restrict ourselves to subjective estimates of certain parameters of the function, to be obtained via questionnaires.

seems reasonable. At this point it should be clear how to estimate the expected utility contribution of a program in a given area i. It is given by

$$(7)\quad w_i E[\Delta f_i] = w_i \{E[f_i(a_i'/a_i^{o}) - f_i(a_i^{o})]$$

$$= w_i \{p_i f_i(a_i^{o}+\delta_i) + (1-p_i) f_i(a_i^{o}) - f_i(a_i^{o})\}$$

$$= w_i\, p_i \left[f_i(a_i^{o}+\delta_i) - f_i(a_i^{o}) \right]$$

and all the values required for this computation are now available.

The choice of δ_i is a tactical matter. It could be a number such that $a_i^{o}+\delta_i$ represents the aspiration level of the decision maker for area i, given an initial percentile score of a_i^{o}. Clearly this aspiration level should be realistic, namely a score which has a good chance of being achieved; $(a_i^{o}+\delta_i)$ may, for instance, also represent one's feelings about the mean, the median or the mode of the posterior probability distribution of scores $g_i(a_i'/a_i^{o})$--illustrating the intimate relationship between δ_i and the probability of success p_i. Alternatively, δ_i may be the smallest number of percentile points such that a score of $a_i^{o}+\delta_i$ would be considered an improvement over a_i^{o} (a subjective assessment must be made in this case as to whether or not the posterior distribution $g_i(a'_i/a_i^{o})$ has really shifted to the right--relative to the prior distribution $g_i(a_i)$--when a_i' takes on the value $a_i^{o}+\delta_i$). A third way may be to simply let δ_i equal a constant, say 5 percentile points, for all subgoal areas. Clearly there must

be tradeoffs between the amount of work which goes into and the information derived from the implementation of each of these three possible methods; however, such tradeoffs have not yet been thoroughly investigated.* At this time the first and third methods are recommended.

*For example it is clear that the first method requires less work than the second; however, the second may provide information about the instantaneous or marginal rate of increase in expected utility associated with a particular program (somewhat in the fashion of a partial derivative). If this second method were applied for several values of δ_i, one may be able to obtain information about the "suspected" decreasing returns, in expected utility, as the "intensity" of a program is increased.

This publication is published pursuant to a contract with the U.S. Office of Education, Department of Health, Education and Welfare. Points of view or opinions stated do not necessarily represent official U.S.O.E. position or policy.